Evidence for \( tt\bar{t}t \) production in the multilepton final state in proton-proton collisions at \( \sqrt{s} = 13 \) TeV with the ATLAS detector

ATLAS Collaboration

DOI
10.1140/epjc/s10052-020-08509-3

Publication date
2020

Document Version
Final published version

Published in
European Physical Journal C

License
CC BY

Citation for published version (APA):
ATLAS Collaboration (2020). Evidence for \( tt\bar{t}t \) production in the multilepton final state in proton-proton collisions at \( \sqrt{s} = 13 \) TeV with the ATLAS detector. European Physical Journal C, 80(11), [1085]. https://doi.org/10.1140/epjc/s10052-020-08509-3

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)
Evidence for $t\bar{t}t\bar{t}$ production in the multilepton final state in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector

ATLAS Collaboration
CERN, 1211 Geneva 23, Switzerland

Received: 30 July 2020 / Accepted: 30 September 2020 / Published online: 24 November 2020
© CERN for the benefit of the ATLAS collaboration 2020

Abstract A search is presented for four-top-quark production using an integrated luminosity of 139 fb$^{-1}$ of proton–proton collision data at a centre-of-mass energy of 13 TeV collected by the ATLAS detector at the LHC. Events are selected if they contain a same-sign lepton pair or at least three leptons (electrons or muons). Jet multiplicity, jet flavour and event kinematics are used to separate signal from the background through a multivariate discriminant, and dedicated control regions are used to constrain the dominant backgrounds. The four-top-quark production cross section is measured to be $(24^{+7}_{-6})$ fb. This corresponds to an observed (expected) significance with respect to the background-only hypothesis of 4.3 (2.4) standard deviations and provides evidence for this process.

Contents
1 Introduction ........................................ 1
2 The ATLAS detector ............................... 2
3 Object and event selection ........................ 3
4 Monte Carlo samples .............................. 4
5 Background estimation ............................ 5
  5.1 Fake/non-prompt lepton background and $t\bar{t}W+$jets production .......................... 5
  5.2 Charge misassignment background ............ 6
6 Signal discrimination .............................. 7
7 Systematic uncertainties ......................... 8
  7.1 Experimental uncertainties ................. 8
  7.2 Signal modelling uncertainties ............ 8
  7.3 Modelling uncertainties in irreducible background ........................ 8
  7.4 Modelling uncertainties in reducible background ........................ 9
8 Results ........................................... 10
9 Conclusion ...................................... 14
References ........................................... 15

1 Introduction

Being the heaviest known elementary particle of the Standard Model (SM), the top quark has a large coupling to the SM Higgs boson and is predicted to have large couplings to hypothetical new particles in many models beyond the SM (BSM). In that respect, rare processes involving the top quark are particularly relevant to study. Among these, the production of four top quarks ($t\bar{t}t\bar{t}$) is predicted by the SM but has not been observed yet. The $t\bar{t}t\bar{t}$ cross section is sensitive to the magnitude and CP properties of the Yukawa coupling of the top quark to the Higgs boson since four top quarks can be produced via an offshell SM Higgs boson [1,2]. Enhancements of the $t\bar{t}t\bar{t}$ cross section ($\sigma_{t\bar{t}t\bar{t}}$) are expected in many BSM scenarios, such as gluino pair production in supersymmetry theories [3,4], pair production of scalar gluons [5,6], or the production of a heavy pseudoscalar or scalar boson in association with a top-quark pair ($t\bar{t}$) in Type II two-Higgs-doublet models (2HDM) [7–9]. Within an effective field theory framework [10], the BSM contribution to $t\bar{t}t\bar{t}$ production can be parameterised by non-renormalisable effective couplings and can be expressed for instance via a $t\bar{t}t\bar{t}$ contact interaction.

The cross section of the SM production of four top quarks from proton–proton ($pp$) collisions at $\sqrt{s} = 13$ TeV is predicted to be $\sigma_{t\bar{t}t\bar{t}} = 12.0$ fb with a relative scale uncertainty of $\pm 20\%$ at next-to-leading order (NLO) in QCD including electroweak corrections [11]. Examples of Feynman diagrams for $t\bar{t}t\bar{t}$ QCD production in the SM are shown in Fig. 1.

In the SM, the top quark is expected to decay into a $W$ boson and a $b$-quark with a branching ratio of approximately 100%. Thus, the $t\bar{t}t\bar{t}$ process will give rise to $W^+W^-W^+W^- bb\bar{b}\bar{b}$ events which then produce different final states depending on the hadronic or leptonic decay mode of the $W$ bosons. This paper considers events that contain exactly two isolated leptons with the same electric charge.
charge (2LSS) or events with at least three isolated leptons\(^1\) (3L), having branching fractions of 7 and 5%, respectively. Although this channel, referred to as 2LSS/3L, has a small branching fraction, it benefits from low levels of background. The \(t\bar{t}\) topology is characterised by high jet and \(b\)-jet multiplicities and high overall energy, which can be quantified as a large value for the scalar sum of the transverse momenta of objects in the event.

A number of SM processes can produce events with topologies similar to those of \(t\bar{t}\) events and thus are backgrounds to \(t\bar{t}\) production. The dominant source is \(t\bar{t}\) production in association with other particles, such as a Higgs boson (\(t\bar{t}H+\text{jets}\)), \(W\) boson (\(t\bar{t}W+\text{jets}\)), or \(Z\) boson (\(t\bar{t}Z+\text{jets}\)). Smaller contributions are expected from \(t\bar{t}WW\), multi-boson production, single-top-quark production, and \(t\bar{t}\) events. Significant backgrounds also come from events where one of the leptons has a misassigned charge and events that contain leptons arising from heavy-flavour decays, photon conversions or misidentified jets, the latter three being collectively referred to as ‘fake/non-prompt’. The heavy-flavour decays are the dominant source for muons, while other sources mostly affect electrons. The charge misassignment and fake/non-prompt background comes mainly from \(t\bar{t}\) events.

In the analysis described in this paper, signal events are separated from background events using a multivariate discriminant. A fit is then performed on the distribution of the multivariate discriminant in the signal-enriched region. Background-enriched regions are also added to the fit to determine the normalisations of the \(t\bar{t}W+\text{jets}\) background and of some sources of fake/non-prompt background.

ATLAS and CMS previously searched for \(t\bar{t}\bar{t}\bar{t}\) production in 13 TeV \(\mathrm{pp}\) collisions. The ATLAS search combined results in the 2LSS/3L channel with those in a channel comprising single-lepton events and dilepton events with two opposite-sign charged leptons (called the 1L/2LOS channel). This analysis used 36 fb\(^{-1}\) of data and led to an observed (expected) significance of 2.8 (1.0) standard deviations \([12,13]\). The CMS combination of the 1L/2LOS and 2LSS/3L channels using 36 fb\(^{-1}\) quotes an observed (expected) significance of 1.4 (1.1) standard deviations \([14]\). The latest CMS search using 137 fb\(^{-1}\) in the 2LSS/3L channel leads to an observed (expected) significance for the \(t\bar{t}\bar{t}\bar{t}\) signal of 2.6 (2.7) standard deviations \([15]\).

### 2 The ATLAS detector

The ATLAS experiment \([16–18]\) at the LHC is a multi-purpose particle detector with a forward–backward symmetric cylindrical geometry and a nearly \(4\pi\) coverage in solid angle.\(^2\) It consists of an inner tracking detector (ID) surrounded by a thin superconducting solenoid providing a 2T axial magnetic field, electromagnetic (EM) and hadron calorimeters, and a muon spectrometer. The inner tracking detector covers the pseudorapidity range \(|\eta| < 2.5\). It consists of silicon pixel, silicon microstrip, and transition radiation tracking detectors. Lead/liquid-argon (LAr) sampling calorimeters provide electromagnetic (EM) energy measurements with high granularity. A steel/scintillator-tile hadron calorimeter covers the central pseudorapidity range \((|\eta| < 1.7\). The endcap and forward regions are instrumented with LAr calorimeters for both EM and hadronic energy measurements up to \(|\eta| = 4.9\). The muon spectrometer (MS) surrounds the calorimeters and is based on three large air-core toroidal superconducting magnets with eight coils each. The field integral of the toroids ranges between 2.0 and 6.0 Tm across most of the detector. The MS includes a system of precision tracking chambers and fast detectors for

---

\(^1\) Throughout the paper, leptons refer to either electrons or muons, which can include those that come from a \(\tau\)-lepton decay.

\(^2\) ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the centre of the detector and the \(z\)-axis along the beam pipe. The \(x\)-axis points from the IP to the centre of the LHC ring, and the \(y\)-axis points upwards. Cylindrical coordinates \((r, \phi)\) are used in the transverse plane, \(\phi\) being the azimuthal angle around the \(z\)-axis. The pseudorapidity is defined in terms of the polar angle \(\theta\) as \(\eta = -\ln \tan(\theta/2)\). Angular distance is measured in units of \(\Delta R \equiv \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2}\).
triggering. A two-level trigger system is used to select events. The first-level trigger is implemented in hardware and uses a subset of the detector information to keep the accepted rate below 100 kHz. This is followed by a software-based trigger that reduces the accepted event rate to 1 kHz, on average, depending on the data-taking conditions [19].

3 Object and event selection

Data used in this analysis were collected by the ATLAS detector between 2015 and 2018 at $\sqrt{s} = 13$ TeV. Only events for which all detector subsystems were operational are considered. The data set corresponds to an integrated luminosity of 139 fb$^{-1}$ [20, 21].

Events were collected using single-lepton or dilepton triggers. Single-lepton triggers select events with leptons satisfying either low transverse momentum ($p_T$) thresholds and an isolation requirement, or a looser identification criterion and higher thresholds with no isolation requirement. The lowest $p_T$ thresholds used in the single-lepton triggers varied from 20 to 26 GeV depending on the lepton flavour and the data-taking period [22, 23]. The $p_T$ thresholds used in the dilepton triggers varied from 8 to 24 GeV depending on the lepton flavour and the data-taking period. Dilepton triggers are used to select events with leptons without requiring any isolation requirement; these are used to validate the fake/non-prompt background estimation.

Events are required to have at least one vertex reconstructed from at least two ID tracks with transverse momenta of $p_T > 0.4$ GeV. The primary vertex for each event is defined as the vertex with the highest sum of $p_T^2$ over all associated ID tracks [24].

Electron candidates are reconstructed from energy deposits in the EM calorimeter associated with ID tracks [25] and are required to have a calorimeter energy cluster with pseudorapidity $|\eta \text{_{cluster}}| < 2.47$, excluding the transition region between the barrel and the endcap calorimeters ($|\eta \text{_{cluster}}| \notin [1.37, 1.52]$). Muon candidates are reconstructed by combining tracks in the ID with tracks in the MS [26] and are required to have $|\eta| < 2.5$. Both the electron and muon candidates are required to have $p_T > 28$ GeV. The transverse impact parameter divided by its estimated uncertainty, $|d_0|/\sigma(d_0)$, is required to be lower than five (three) for electron (muon) candidates. The longitudinal impact parameter must satisfy $|z_0 \sin(\theta)| < 0.5$ mm for both lepton flavours. Electrons are required to meet the ‘Tight’ likelihood-based identification criterion and to be isolated using criteria based on the properties of the topological clusters in the calorimeter and of the ID tracks around the reconstructed electron [25].

Muons are required to meet the ‘Medium’ cut-based identification criterion, which includes requirements on the number of hits in the ID and MS as well as requiring compatibility between momentum measurements in the ID and MS. Muons also have to satisfy the isolation requirement based on the properties of ID tracks around the reconstructed muon.

To reduce the impact of charge misassignment background, an additional requirement is imposed on electrons in the $e^\pm e^\pm$ and $e^\pm \mu^\pm$ channels. This requirement is based on the score of a boosted decision tree (BDT) that uses the calorimeter cluster and track properties of the electron [25] and is trained on data enriched in $Z \rightarrow ee$ events to separate events with correct and incorrect electron charge assignments. The chosen requirement on the BDT score removes approximately 90% of electrons with a wrong charge assignment while selecting 98% of electrons with correctly measured charge.

Jets are reconstructed from topological clusters [27] of energy deposits in the calorimeters using the anti-$k_t$ algorithm [28, 29] with a radius parameter of $R = 0.4$ and are calibrated as described in Ref. [30]. Jets are required to have $p_T > 25$ GeV and $|\eta| < 2.5$. To reduce the effect from additional $pp$ collisions in the same or a nearby bunch crossing, collectively referred to as pile-up, jets with $p_T < 120$ GeV and $|\eta| < 2.4$ are considered only when they satisfy a requirement based on the output of a multivariate classifier called the jet-vertex-tagger (JVT) [31]. Events that contain at least one jet arising from non-collision sources or detector noise are rejected by a set of quality criteria [32]. The MV2c10 multivariate algorithm [33] is used to identify jets containing $b$-hadrons. A jet is considered $b$-tagged if it passes the operating point corresponding to 77% average efficiency for $b$-quark jets in simulated $t\bar{t}$ events with the corresponding rejection factors against light-quark/gluon jets and $c$-quark jets of 110 and 4, respectively.

A sequential overlap removal procedure is applied to avoid the same calorimeter energy deposit or the same track being reconstructed as two different objects. As a first step, electrons sharing their track with a muon candidate are removed. Next, the closest jet within $\Delta R = \sqrt{(\Delta \eta)^2 +(\Delta \phi)^2}$ = 0.2 of an electron is removed. Then, electrons within $\Delta R = 0.4$ of a remaining jet are removed since they likely arise from $b$- or $c$-decays. After that, jets with fewer than three associated tracks that are within $\Delta R = 0.2$ of a muon are removed. Finally, muons are removed if their tracks are within $\Delta R = 0.4 + 10 \text{GeV} / p_T$ of any remaining jets as they also likely arise from $b$- or $c$-decays.

The missing transverse momentum in the event, whose magnitude is denoted in the following by $E_T^{\text{mis}}$, is defined as the negative vector sum of the $p_T$ of the reconstructed and calibrated objects in the event [34]. This sum includes the momenta of the ID tracks that are matched to the primary vertex and are not associated with any other objects.

$$ |\rho| = \frac{1}{2} \ln \frac{E + p_T}{E - p_T}, \text{ where } E \text{ is the energy and } p_T \text{ is the component of the momentum along the beam pipe.}$$
The events are required to have one same-sign lepton pair or at least three leptons without charge requirement. Each event must have at least one reconstructed lepton that matches a lepton that fired the trigger. Events with two same-sign electrons are required to have the invariant mass $m_{ee} > 15$ GeV and $|m_{ee} - 91$ GeV$| > 10$ GeV to reduce the charge misassignment background coming from low-mass resonances and Z-boson decay. In events with at least three leptons, all the opposite-sign same-flavour lepton pairs are required to satisfy $|m_{ll} - 91$ GeV$| > 10$ GeV to reduce the contamination from Z-boson decay.

Events arising from $t\bar{t}t\bar{t}$ production are selected by exploiting the high multiplicities of light-flavour jets and $b$-tagged jets as well as the large overall event activity. This last property is probed by the scalar sum of the transverse momenta of the isolated leptons and jets in the event, denoted by $H_T$. The inclusive signal region (SR) is defined by requiring at least six jets, at least two $b$-tagged jets, and $H_T$ above 500 GeV.

4 Monte Carlo samples

Production of $t\bar{t}t\bar{t}$ events is modelled according to the SM expectation. The nominal sample used to model the $t\bar{t}t\bar{t}$ signal was generated using the MadGraph5_aMC@NLO v2.6.2 [35] generator, which produces matrix elements (ME) at NLO in the strong coupling constant $\alpha_S$, with the NNPDF3.1nlo [36] PDF set. The functional form of the renormalisation and factorisation scales was set to $\mu_F = \mu_T = m_T/4$, where $m_T$ is defined as the scalar sum of the transverse masses $\sqrt{p_T^2 + m^2}$ of the particles generated from the matrix element calculation, following Ref. [11]. The parton shower, fragmentation, and underlying event were simulated using Pythia 8.230 [37] with the A14 set of tuned parameters (tune) [38] and the NNPDF2.3lo PDF set. The top-quark mass $m_{top}$ in this sample and in all other simulated samples is set to 172.5 GeV. An alternative $t\bar{t}t\bar{t}$ sample generated with the same MadGraph5_aMC@NLO set-up but interfaced to Herwig 7.04 [39,40] with the H7UE tune [40] and the NNHTT2014 LO PDF set [41] is used to evaluate uncertainties due to the choice of parton shower and hadronisation model. In order to mitigate the effect of the large fraction of negative weights present in the nominal sample that would be detrimental for training of the multivariate discriminant used to separate signal from background (see Sect. 6), an additional sample with settings similar to the nominal ones was generated using leading-order (LO) matrix elements. Good agreement between the distributions of the kinematic variables used by the multivariate discriminant simulated at LO and NLO was observed.

The $t\bar{t}W$ simulated events were generated using the Sherpa 2.2.1 [42] generator with the NNPDF3.0nlo PDF set and the tune provided by the Sherpa authors. The ME was calculated for up to one additional parton at NLO QCD and up to two partons using the five-flavour scheme, including $c$- and $b$-quarks, at LO QCD using the Comix [42] and Open-Loops [43,44] libraries, and was merged with the Sherpa parton shower [45] using the MEPS@NLO prescription [46–49] and a merging scale of 30 GeV. The renormalisation and factorisation scales were set to $\mu_F = \mu_T = m_T/2$. The simulated $t\bar{t}W$ sample is normalised to the cross section of 601 fb computed at NLO in QCD with the leading NLO electroweak corrections [50–52]. An alternative $t\bar{t}W$ sample was generated at NLO in QCD with no additional partons using the MadGraph5_aMC@NLO v2.3.3 generator with the same PDF as the nominal sample. The events were interfaced to Pythia 8.210 using the A14 tune and the NNPDF2.3lo PDF set.

The production of $t\bar{t}(Z/\gamma^*)$ events was modelled at NLO in QCD using the MadGraph5_aMC@NLO v2.3.3 generator in the five-flavour scheme with the NNPDF3.0nlo PDF set interfaced to Pythia 8.230 using the A14 tune and the NNPDF2.3lo PDF set. It is normalised to the inclusive $t\bar{t}e^+e^−$ cross section of 880 fb computed at NLO in QCD [50–52], including off-shell Z and $\gamma^*$ contributions with $m(e^+e^-) > 5$ GeV. An alternative sample was generated with NLO matrix elements using Sherpa 2.2.1 and the same PDF as the nominal sample.

The production of $t\bar{t}$ and single-top-quark events was modelled using the Powheg-Box [53–56] v2 generator at NLO in QCD with the NNPDF3.0nlo PDF set. In the $t\bar{t}$ sample the $h_{\text{damp}}$ parameter$^4$ was set to 1.5 $m_{top}$ [57]. The overlap between the $t\bar{t}$ and the $tW$ final states was removed using the diagram removal technique [58]. The $t\bar{t}$ and single-top-quark simulated samples are normalised to the cross sections calculated at next-to-next-to-leading order (NNLO) in QCD including the resummation of next-to-next-to-leading logarithmic (N NLL) soft-gluon terms [59–62].

The production of $t\bar{t}H$ events was modelled by the Powheg-Box v2 generator at NLO in QCD using the five-flavour scheme with the NNPDF3.0nlo PDF set, with the $h_{\text{damp}}$ parameter set to $1.5 \times (2m_{top} + m_H)/2$ where the Higgs boson mass is $m_H = 125$ GeV. The simulated sample is normalised to the cross section computed at NLO in QCD with the leading NLO electroweak corrections [50–52]. An alternative sample generated using the MadGraph5_aMC@NLO v2.3.3 generator with the same PDF as the nominal sample was used.

$^4$ The $h_{\text{damp}}$ parameter controls the transverse momentum $p_T$ of the first additional emission beyond the leading-order Feynman diagram in the parton shower and therefore regulates the high-$p_T$ emission against which the $t\bar{t}$ system recoils.
settings is used to evaluate the uncertainty in $t\bar{t}H$ modelling due to the generator choice.

The $t\bar{t}WZ$ events were generated at NLO in QCD using the MadGraph5_aMC@NLO v2.3.3 generator with the NNPDF3.0nlo PDF set. The other rare top-quark processes, namely the production of $t\bar{t}$, $t\bar{t}WW$, $t\bar{t}ZZ$, $t\bar{t}WZ$, $t\bar{t}HH$ and $t\bar{t}WH$, and $t\bar{t}t$, were modelled using the MadGraph5_aMC@NLO generator at LO in QCD. The $t\bar{t}$, single-top-quark, $t\bar{t}H$, $t\bar{t}Z$, $t\bar{t}W$, $t\bar{t}WW$, $t\bar{t}ZZ$, $t\bar{t}WZ$, $t\bar{t}HH$, $t\bar{t}WH$ and $t\bar{t}t$ generated samples were all interfaced with Pythia 8.230 using the A14 tune and the NNPDF2.3lo PDF set. Rare top-quark background contributions are normalised using their NLO QCD theoretical cross sections. The reducible backgrounds originate mainly from diboson or triboson production, $VH$ production in association with jets, and rare processes ($t\bar{t}WW$, $t\bar{t}Z$, $t\bar{t}Zq$, $t\bar{t}t$). The irreducible background is evaluated using MC simulation normalised to the SM cross sections, except $t\bar{t}W$+jets for which the normalisation is corrected using data in a dedicated control region. The different treatment for the $t\bar{t}W$+jets background is motivated by theoretical studies [69] showing that electroweak corrections not included in the simulation have a significant effect as well as by the large $t\bar{t}W$+jets background normalisation factor found in recent measurements in similar phase space [70].

The reducible backgrounds originate mainly from $t\bar{t}$+jets and $t\bar{t}W$+jets production and have prompt leptons with misassigned charge (Q mis-id) or fake/non-prompt leptons. This fake/non-prompt background, together with $t\bar{t}W$+jets, is evaluated using the template method (cf. Sect. 5.1). The charge misassignment background is defined for the 2LSS only. It arises mainly from $t\bar{t}$+jets events with an opposite-charge lepton pair in which the charge of one electron is mismeasured either due to bremsstrahlung photon emission followed by its conversion ($e^{\pm} \rightarrow e^{\pm}\gamma \rightarrow e^{\pm}e^{\mp}e^{-}$) or due to mismeasured track curvature. This background is evaluated using a data-driven method (cf. Sect. 5.2). The charge misassignment rate is negligible for muons due to the low probability of bremsstrahlung and the large lever arm of the muon spectrometer.

The estimated yield from each source of background is given in Sect. 8.

5.1 Fake/non-prompt lepton background and $t\bar{t}W$+jets production

The template method used to estimate the fake/non-prompt background relies on the simulation to model the kinematic distributions of background processes arising from fake and non-prompt leptons and on control regions to determine their normalisations. These control regions are included in the fit together with the signal region, and the normalisation factors are determined simultaneously with the $t\bar{t}$+jets signal.

The following main contributions of the fake/non-prompt background are distinguished:

- events with one non-prompt electron (muon) from heavy-flavour decay, $HF e$ ($HF \mu$),
- events with one non-prompt electron originating from photon conversion taking place in the detector material (Mat. Conv.),
- events with a virtual photon ($\gamma^*$) leading to an $e^+e^-$ pair (Low $m_{\gamma^*}$).
Table 1 Summary of the signal and control regions used in the template fit. The variable $m_{ee}^{CV}$ ($m_{ee}^{PV}$) is defined as the invariant mass of the system formed by the track associated with the electron and the closest track at the conversion (primary) vertex. $N_j$ ($N_b$) indicates the jet (b-tagged jet) multiplicity in the event. $H_T$ is defined as the scalar sum of the transverse momenta of the isolated leptons and jets.

<table>
<thead>
<tr>
<th>Region</th>
<th>Channel</th>
<th>$N_j$</th>
<th>$N_b$</th>
<th>Other requirements</th>
<th>Fitted variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>SR</td>
<td>2LSS/3L</td>
<td>$\geq 6$</td>
<td>$\geq 2$</td>
<td>$H_T &gt; 500$</td>
<td>BDT</td>
</tr>
<tr>
<td>CR Conv.</td>
<td>$e^\pm e^\pm</td>
<td>e^\pm \mu^\pm$</td>
<td>$4 \leq N_j &lt; 6$</td>
<td>$1$</td>
<td>$m_{ee}^{CV} \in [0, 0.1 \text{ GeV}]$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>$200 &lt; H_T &lt; 500 \text{ GeV}$</td>
<td></td>
</tr>
<tr>
<td>CR HF e</td>
<td>$eee \</td>
<td>ee \mu$</td>
<td>$- \</td>
<td>-$</td>
<td>$= 1$</td>
</tr>
<tr>
<td>CR HF $\mu$</td>
<td>$e\mu</td>
<td>\mu\mu$</td>
<td>$- \</td>
<td>-$</td>
<td>$= 1$</td>
</tr>
<tr>
<td>CR $tW$</td>
<td>$e\pm \mu^\pm</td>
<td>\mu^\pm\mu^\pm$</td>
<td>$\geq 4$</td>
<td>$\geq 2$</td>
<td>$m_{ee}^{CV} \notin [0, 0.1 \text{ GeV}],</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>For $N_b = 2$, $H_T &lt; 500 \text{ GeV}$ or $N_j &lt; 6$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>For $N_b \geq 3$, $H_T &lt; 500 \text{ GeV}$</td>
<td></td>
</tr>
</tbody>
</table>

The minor components of the fake/non-prompt background arising from events with a lepton originating from light-meson decay (LF) or with a jet misidentified as a lepton (other fakes) are determined from MC simulation.

Several control regions, non-overlapping with the signal region, are defined to determine the normalisation of various components of the fake/non-prompt background from data. Each region is required to have a dominant component or a variable with good discriminating power between different components. Since events arising from $t\bar{t}W$+jets production represent a large contribution in all control and signal regions, the normalisation of that process is also determined using a dedicated control region. In total, four control regions with their corresponding discriminating variables are used in the analysis. They are summarised in Table 1 and are defined below:

- ‘CR Conv.’ is enriched in background events arising from both material photon conversion and processes with a virtual photon leading to an $e^+e^-$ pair. For each electron in the selected $e^\pm e^\pm$ or $e^\pm \mu^\pm$ events, the invariant mass of the system formed by the track associated with the electron and the closest track at the conversion (primary) vertex $m_{ee}^{CV}$ ($m_{ee}^{PV}$) is computed. The conversion vertex is defined as the point where the track from the electron and its closest track in $\Delta R$ have the same $\phi$. The control region is then obtained by selecting events with at least four or five jets, at least one identified b-jet, with low $m_{ee}^{CV}$ and using the $m_{ee}^{PV}$ distribution in the fit to separate the material conversion and the $\gamma^*$ components from each other. Virtual photons lead to a lepton pair originating from the primary vertex, having a low $m_{ee}^{PV} \sim m_{\gamma^*}$ and a low conversion radius. Material conversions happen further away from the primary vertex with a larger conversion radius, and the track extrapolation induces a larger apparent invariant mass. According to the MC simulation, the background arising from both $\gamma^*$ and material conversions accounts for around 40% of the total event yield in this control region.

- ‘CR HF e’ (‘CR HF $\mu$’) is enriched in background events with an electron (muon) from heavy-flavour decay. This region is defined by selecting events with three leptons, namely $eee$ and $ee\mu$ ($\mu\mu\mu$ and $\mu\mu\mu$) for CR HF e (CR HF $\mu$), and exactly one identified b-jet. This selection targets $t\bar{t}$ dileptonic decays with an extra non-prompt lepton in events with low $H_T$. The number of events in the region is obtained using the maximum-likelihood fit. According to the MC simulation, the background with an electron (muon) coming from heavy-flavour decay accounts for around 40% (50%) of the total event yield in the CR HF e (CR HF $\mu$).

- ‘CR $tW$’ is enriched in $t\bar{t}W$+jets events. This region is obtained by selecting $e\mu$ and $\mu\mu$ events with at least four jets and two b-jets which are neither in other CRs nor in the SR. Events containing electrons with $|\eta| > 1.5$ and $ee$ final states are not considered, in order to reduce the contamination arising from charge misassignment background. The sum of the lepton $p_T$ provides discrimination from other processes and is used in the maximum-likelihood fit. According to the MC simulation, the $t\bar{t}W$+jets background accounts for around 33% of the total event yield in this control region.

5.2 Charge misassignment background

The probability for an electron to have its charge incorrectly assigned is measured using a data sample of $Z \rightarrow ee$ events requiring the invariant mass of the electron pair to be within 10 GeV of the Z-boson mass and without any requirement on the charge of the two electron tracks. The background contamination is subtracted using a sideband method [12]. The charge misassignment rate is parameterised as a function of electron $p_T$ and $|\eta|$, except for the conversion control region defined in Sect. 5.1, where it is also parameterised as a function of the invariant mass of the electron track and its closest
track assuming that both tracks originate from the primary vertex. The charge misassignment rate varies from 0.002 to 4% depending on the electron \( p_T \) and \( |\eta| \).

The expected number of events arising from charge misassignment background is determined by applying the measured charge misassignment rate to data events satisfying the requirements of the kinematic selection of the ZLLS channel, except that the two leptons are required to be of opposite charge. In this sample, each event is weighted according to the value of the charge misassignment rate of each electron in the event.

6 Signal discrimination

The background composition of the SR is largely dominated by the production of top-quark pairs in association with additional jets and/or bosons. To separate signal from background events, a multivariate discriminant is built in the SR by combining several input observables into a boosted decision tree (BDT). This set of variables and the BDT hyperparameters are optimised to maximise the integral under the receiver operating characteristic (ROC) curve of the BDT. In total, 12 observables are selected, based on their discrimination power and the requirement of good modelling. Among them, the pseudo-continuous \( b \)-tagging discriminant score [33] summed over all the jets in the event is the best discriminating variable due to the four \( b \)-jets being produced mainly in signal events. The pseudo-continuous \( b \)-tagging discriminant score is an integer from 1 to 5 assigned to a jet, based on the operating point of the \( b \)-tagging algorithm it passes, with a value of five corresponding to the most \( b \)-like jet. For the signal region selection the minimum score is 10 for an event with exactly two \( b \)-tagged jets out of exactly six jets. The minimum distance \( \Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2} \) between two leptons among all possible pairs is the second best discriminating variable since it provides good discrimination for events with at least three leptons. The distributions for those two variables are shown in Fig. 2. The other input variables are the leading lepton \( p_T \), \( E_T^{\text{miss}} \), the \( p_T \) of the leading and second-leading jets, the \( p_T \) of the sixth jet, the \( p_T \) of the leading \( b \)-jet, the scalar sum of transverse momenta over all leptons and all jets excluding the leading \( p_T \) jet, the sum of distances \( \Delta R \) between two leptons for all possible pairs, the maximum distance \( \Delta R \) between a \( b \)-jet and a lepton among all possible pairs, and the minimum distance \( \Delta R \) between a jet and a \( b \)-jet among all possible pairs. Taking into account all uncertainties, no significant discrepancy between data and predicted background was found for these variables in the various CRs.

The BDT training is performed inclusively, both in lepton flavour and lepton multiplicity for events passing the SR requirements. The LO \( tt\bar{t} \) simulated signal sample is used...
in the training. The background sample corresponds to the total expected background, as predicted by the simulation.

7 Systematic uncertainties

Various sources of systematic uncertainty impact the estimated signal and background rates, the migration of events between regions, and the shape of the fitted discriminant distributions. They can be classified into the experimental uncertainties and modelling uncertainties of the $t\bar{t}t\bar{t}$ signal and of backgrounds, as described below. The impact of each source of systematic uncertainty on the final result can be found in Sect. 8.

7.1 Experimental uncertainties

The uncertainty in the combined 2015–2018 integrated luminosity is 1.7% [20], obtained using the LUCID-2 detector [21] for the primary luminosity measurements. To account for the difference between the pile-up distributions in data and MC simulations, an uncertainty related to the scale factors used to adjust the MC pile-up to the data pile-up profile is applied.

For electrons and muons, the reconstruction, identification, isolation and trigger performances as well as the lepton momentum scale and resolution differ between data and MC simulation. To correct for these differences, scale factors for each are applied. These scale factors were estimated using the tag-and-probe method [25, 26], which is performed using the leptonic decays of $Z$ and $W$ bosons and of $J/\psi$ mesons. The associated systematic uncertainties are then propagated to the final distributions used in this analysis.

To determine the jet energy scale (JES) and its associated uncertainty, information from test-beam data, LHC collision data, and simulation was used, as described in Ref. [71]. The JES uncertainty is decomposed into a set of 30 uncorrelated components, 29 of which are used per event depending on the type of simulation. The jet energy resolution (JER) is measured separately for data and MC simulation using in situ techniques, similar to those in Ref. [72]. Its uncertainty is represented by nine components accounting for jet-$p_T$ and $\eta$-dependent differences between simulation and data, eight of which are used per event depending on the type of simulation. The systematic uncertainty associated with the JVT is obtained by varying the scale factor used to correct the JVT efficiency in simulation up and down within its uncertainties [31].

The $b$-tagging efficiencies and mistagging rate are measured in data using the same methods as are described in Refs. [33, 73, 74], with the systematic uncertainties due to $b$-tagging efficiency and the mistagging rates calculated separately. The impact of the uncertainties on the $b$-tagging calibration is evaluated separately for $b$-jets, $c$-jets and light-flavour jets in the MC samples.

The $E_T^{\text{miss}}$ uncertainty due to a possible miscalibration of its soft-track component is derived from data–MC comparisons of the $p_T$ balance between the hard and soft $E_T^{\text{miss}}$ components [34].

7.2 Signal modelling uncertainties

Several sources of modelling uncertainty are considered for the $t\bar{t}t\bar{t}$ signal. The uncertainty due to missing higher-order QCD corrections is determined by varying the renormalisation and the factorisation scales simultaneously by factors of 2.0 and 0.5 relative to the central value. The uncertainty related to the choice of parton shower and hadronisation model is estimated by comparing the nominal prediction with that obtained using an alternative sample generated with MadGraph5_aMC@NLO interfaced to Herwig 7. The effect of the PDF uncertainty on the signal MC prediction is calculated as the RMS of the signals from the 100 replicas of the NNPDF30_nlo_as_0118 PDF set following the PDF4LHC prescription [75]. Shape and normalisation variations due to the PDF uncertainty are found to be negligible.

7.3 Modelling uncertainties in irreducible background

Modelling uncertainties for the $t\bar{t}W$+jets, $t\bar{t}Z$+jets and $t\bar{t}H$+jets processes are evaluated in a similar way and include the uncertainty due to missing higher-order QCD corrections determined by varying the renormalisation and the factorisation scales simultaneously by factors of 2.0 and 0.5 relative to the central value, and a comparison with alternative generators.

For $t\bar{t}Z$+jets, the nominal MC prediction is compared with an NLO SHERPA sample, while for $t\bar{t}H$+jets the nominal simulation is compared with an NLO MadGraph5_aMC@NLO sample, both described in Sect. 4. A 1% uncertainty from the PDF is assigned to both the $t\bar{t}Z$ and $t\bar{t}H$ processes following the same procedure as described in Sect. 7.2. For $t\bar{t}W$+jets, uncertainties associated with the modelling of additional QCD radiation, with the choice of the ME generator and parton shower, are estimated by comparing the nominal prediction with that of an alternative sample that was generated at NLO with no additional partons using the MadGraph5_aMC@NLO generator with the same scale choice and PDF set as for the nominal sample (cf. Sect. 4).

An uncertainty of 15% (20%) is applied to the $t\bar{t}Z$ ($t\bar{t}H$) total cross section [52, 76]. Since the $t\bar{t}W$+jets normalisation is determined from the fit to data, no cross-section uncertainty is applied to this process. An additional 125% (300%) uncertainty is added for $t\bar{t}W$ production with seven (eight or more) jets. These values correspond to the difference between data and prediction in a $t\bar{t}W$+jets validation region (cf. Sect. 8).
where a data excess is observed for high jet multiplicities. Since the jet multiplicity distribution in a $t\bar{t}Z$+jets validation region shows good agreement between data and prediction, such uncertainty is not considered for $t\bar{t}Z$+jets or for $t\bar{t}H$+jets production due to similarity of their simulation.

The $t\bar{t}W$+jets, $t\bar{t}Z$+jets and $t\bar{t}H$+jets background processes enter the $t\bar{t}t\bar{t}$ signal region if they have additional heavy-flavour jets. Such processes are difficult to model with the MC simulation. To account for this, an uncertainty of 50% is assigned to the events with three generator-level (‘true’) $b$-jets and a separate 50% uncertainty to the events with four or more true $b$-jets. These estimates are based on the measurement of $t\bar{t}$ production with additional heavy-flavour jets [77] and on comparisons between data and prediction in $t\bar{t}\gamma$ events with three and four $b$-tagged jets. They are treated as uncorrelated between the three backgrounds due to the different MC setups used to simulate the $t\bar{t}W$+jets, $t\bar{t}Z$+jets and $t\bar{t}H$+jets backgrounds.

The $t\bar{t}t$ events have similar kinematics to the $t\bar{t}t\bar{t}$ signal, although the rate is expected to be much smaller. However, it is currently unexplored experimentally. Thus a large ad hoc uncertainty of 100% is assigned to its cross section and an additional 50% uncertainty is applied to $t\bar{t}t$ events with four true $b$-jets.

The uncertainty in the $tZ$ and $tWZ$ single-top-quark cross sections is set to 30% [78,79] and that for the $t\bar{t}WW$, $t\bar{t}ZZ$, $t\bar{t}WZ$, $t\bar{t}HH$ and $t\bar{t}WH$ cross sections to 50% [12]. The uncertainty in diboson production is set to 40%, based on studies of the $WZ + b$ process. For each of the other small background processes a large ad hoc cross-section uncertainty of 50% is applied. For all small backgrounds except $t\bar{t}t$ an additional 50% uncertainty is assigned to the events with three true $b$-jets and separately a 50% uncertainty for events with four or more true $b$-jets.

7.4 Modelling uncertainties in reducible background

Uncertainties in the charge misassignment background arise from the following contributions: the statistical uncertainty of the fit to data used to determine the rates; the rate variation due to variation of the dielectron invariant mass requirement; and the rate variation due to a difference between the observed and the predicted misidentification rates when the method is applied to MC simulated events. This uncertainty is determined separately for the material conversion control region, for the $t\bar{t}W$+jets, and for all other control regions, and it is treated as correlated between the regions.

Since the overall normalisations of the material conversion and the virtual photon backgrounds are free parameters in the fit, their uncertainty comes only from the shape of the distributions used in the template fit (cf. Sect. 5.1). For each of these sources, the uncertainty is obtained by comparing data with the Powheg + Pythia8 simulation of $Z(\rightarrow \mu\mu) + \gamma$ and $Z(\rightarrow \mu\mu) + \text{jets}$ production in a region enriched in $Z(\rightarrow \mu\mu) + \gamma$ events. An uncertainty of 25% is applied to the material conversion and to the virtual photon background events fulfilling $m_{ee} > 0.1$ GeV in all control and signal regions to cover the extrapolation from the ‘CR Conv.’ region with $0 < m_{ee} < 0.1$ GeV to the regions with events with larger $m_{ee}$.

The uncertainty in the shape of the distributions of the heavy-flavour non-prompt lepton background is estimated by comparing data with the background prediction, normalised to data, for a loose lepton selection with the isolation requirements dropped and the identification criteria relaxed. The shape uncertainty is derived for each region included in the fit, but these variations are treated as correlated between regions since the physics origin of the uncertainty is common to all of them. This systematic uncertainty is derived separately for electrons and muons.

A normalisation uncertainty of 100% is assigned to the background arising from light-flavour non-prompt leptons. This uncertainty was found to cover any difference between data and prediction in loose lepton regions [70]. An ad hoc uncertainty of 30% is applied to the normalisation of the background arising from the other minor sources of non-prompt leptons from $t\bar{t}$ production. No uncertainty in the shape of the distributions of these backgrounds is considered since their contribution is very small.

Fig. 3 Comparison between data and prediction after the fit (‘Post-Fit’) for the distribution of the BDT score in the SR. The band includes the total uncertainty of the post-fit computation. The ratio of the data to the total post-fit computation is shown in the lower panel. See Sect. 5 for the definitions of the different background categories.
Since the main source of reducible background is \( t\bar{t}+\)jets production, the systematic uncertainty in the modelling of its heavy-flavour content can affect the shape of the template distributions used in the fit. To account for this effect an uncertainty of 30%, based on the measurement of \( t\bar{t} \) production with additional heavy-flavour jets [77], is assigned to the events with three true \( b \)-jets and a separate 30% uncertainty to events with four or more true \( b \)-jets. A small contribution to the reducible background from \( V+\)jets production is determined from simulation and a normalisation uncertainty of 30% is assigned to this background.

### 8 Results

The \( t\bar{t}t\bar{t} \) production cross section and the normalisation factors of the backgrounds are determined via a binned likelihood fit to the BDT score distribution in the signal region and to the yields, or to the discriminating variable distributions, in the four control regions as listed in Table 1. The systematic uncertainties in both the signal and background predictions are included as nuisance parameters in the likelihood function. The maximum-likelihood fit is performed using the RooFit package [80] based on a likelihood function built on the Poisson probability that the observed data are compatible with the model prediction. The value of each nuisance parameter is constrained by a penalty factor present in the likelihood function, while all normalisation factors are unconstrained.

The fit determines the best value of the signal strength \( \mu \), defined as a ratio of the \( t\bar{t}t\bar{t} \) cross section to the SM expectation, its uncertainty, and five normalisation factors: \( \text{NF}_{HF\ e} \) (\( \text{NF}_{HF\ \mu} \)) for the non-prompt electron (muon) background from heavy-flavour decays, \( \text{NF}_{\text{Mat. Conv.}} \) for the background from detector material conversions, \( \text{NF}_{\text{Low } m_{\gamma\gamma}} \) for the contribution of low-mass electron pairs, and \( \text{NF}_{t\bar{t}W} \) for the \( t\bar{t}W+\)jets contribution. For each free parameter, the uncertainty is derived following the asymptotic approximation [81]. An uncertainty of 20% is assigned to the \( t\bar{t}t\bar{t} \) cross section predicted by the SM. The prediction corresponding to all parameters maximising the full likelihood is referred to as the post-fit model.

The best-fit value of \( \mu \) is:

\[
\mu = 2.0 \pm 0.4 \text{(stat)}^{+0.7}_{-0.4} \text{(syst)} = 2.0^{+0.8}_{-0.6}.
\]

### Table 2: Normalisation factors for various backgrounds determined from the fit to the control regions. The uncertainties include both the statistical and systematic uncertainties.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>( \text{NF}_{t\bar{t}W} )</th>
<th>( \text{NF}_{\text{Mat. Conv.}} )</th>
<th>( \text{NF}<em>{\text{Low } m</em>{\gamma\gamma}} )</th>
<th>( \text{NF}_{HF\ e} )</th>
<th>( \text{NF}_{HF\ \mu} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>1.6 ± 0.3</td>
<td>1.6 ± 0.5</td>
<td>0.9 ± 0.4</td>
<td>0.8 ± 0.4</td>
<td>1.0 ± 0.4</td>
</tr>
</tbody>
</table>

### Table 3: Post-fit background and signal yields in the full signal region as well as for events in which the BDT score is also greater than zero. The total systematic uncertainty differs from the sum in quadrature of the different uncertainties due to correlations. Q mis-id refers to the charge misassignment background. Mat. Conv. and Low \( m_{\gamma\gamma} \) refer respectively to events with one non-prompt electron originating from photon conversion in the detector material and to events with a virtual photon leading to an \( e^+e^- \) pair. HF e (HF \( \mu \)) refers to events with one non-prompt electron (muon) from heavy-flavour hadron decay. LF refers to events with a lepton originating from light-meson decay, and ‘Other \( t\bar{t}X \)’ includes events coming from \( t\bar{t}WZ, t\bar{t}ZZ, t\bar{t}W H, t\bar{t}HH \).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>SR</th>
<th>SR and BDT &gt; 0</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t\bar{t}W+)jets</td>
<td>102 ± 26</td>
<td>23 ± 10</td>
</tr>
<tr>
<td>( t\bar{t}WW )</td>
<td>7 ± 4</td>
<td>2 ± 1</td>
</tr>
<tr>
<td>( t\bar{t}Z+)jets</td>
<td>48 ± 9</td>
<td>9 ± 2</td>
</tr>
<tr>
<td>( t\bar{t}H+)jets</td>
<td>38 ± 9</td>
<td>8 ± 2</td>
</tr>
<tr>
<td>Q mis-id</td>
<td>16 ± 1</td>
<td>2.7 ± 0.2</td>
</tr>
<tr>
<td>Mat. Conv.</td>
<td>19 ± 6</td>
<td>3 ± 1</td>
</tr>
<tr>
<td>Low ( m_{\gamma\gamma} )</td>
<td>9 ± 4</td>
<td>0.9 ± 0.5</td>
</tr>
<tr>
<td>HF e</td>
<td>3 ± 3</td>
<td>1 ± 1</td>
</tr>
<tr>
<td>HF ( \mu )</td>
<td>12 ± 6</td>
<td>3 ± 2</td>
</tr>
<tr>
<td>LF</td>
<td>4 ± 5</td>
<td>1 ± 1</td>
</tr>
<tr>
<td>Other fake</td>
<td>6 ± 2</td>
<td>2 ± 1</td>
</tr>
<tr>
<td>VR, VVVV, VH</td>
<td>3 ± 2</td>
<td>0.2 ± 0.2</td>
</tr>
<tr>
<td>( tZZ, tWZ )</td>
<td>5 ± 2</td>
<td>1.0 ± 0.4</td>
</tr>
<tr>
<td>Other ( t\bar{t}X )</td>
<td>3 ± 2</td>
<td>1 ± 1</td>
</tr>
<tr>
<td>( t\bar{t}t )</td>
<td>3 ± 3</td>
<td>2 ± 2</td>
</tr>
<tr>
<td>Total bkg</td>
<td>278 ± 22</td>
<td>59 ± 10</td>
</tr>
<tr>
<td>( t\bar{t}t )</td>
<td>60 ± 17</td>
<td>44 ± 12</td>
</tr>
<tr>
<td>Total</td>
<td>337 ± 18</td>
<td>103 ± 10</td>
</tr>
<tr>
<td>Data</td>
<td>330</td>
<td>105</td>
</tr>
</tbody>
</table>

The systematic uncertainty is determined by subtracting in quadrature the statistical uncertainty, obtained from a fit where all NPs are fixed to their post-fit values, from the total uncertainty. The measured \( \mu \) value is consistent within 1.7 standard deviations with the SM prediction corresponding to \( \mu = 1 \). The probability for the background-only hypothesis to result in a signal-like excess at least as large as seen in data is derived using the profile-likelihood ratio following the procedure described in Ref. [81]. From this, the significance of the observed signal is found to be 4.3 standard deviations, while 2.4 standard deviations are expected. Figure 3 shows the distribution of the BDT score in the signal region after
Fig. 4 Comparison between data and prediction after the fit ('Post-Fit') for the yields or distributions of the discriminating variables used in the fit in each CR (see Table 1). The band includes the total uncertainty of the post-fit computation. The ratio of the data to the total post-fit computation is shown in the lower panel. The first and last bins contain underflow and overflow events, respectively. See Sect. 5 for the definitions of the different background categories.

performing the fit. Good agreement is observed between data and the fitted prediction.

The fitted signal strength is converted into an inclusive cross section using the SM \( \bar{t}t \bar{t}t \) predicted cross section of \( \sigma_{\bar{t}t \bar{t}t} = 12.0 \pm 2.4 \) fb computed at NLO in QCD and electroweak couplings [11] and excluding its uncertainty. The measured \( \bar{t}t \bar{t}t \) production cross section is then:

\[ \sigma_{\bar{t}t \bar{t}t} = 24 \pm 5 \text{(stat)} \pm 4 \text{(syst)} \text{ fb} = 24 \pm 7 \text{ fb}. \]

The normalisation factors of the different background sources determined from the fit are shown in Table 2. The post-fit background and signal yields are shown in Table 3.
The normalisation factor for the \( ttW \)+jets background is compatible with the observation from the previous ATLAS \( ttH \) search [70] where the reference theoretical \( ttW \)+jets background cross section was scaled up by 20% to account for extra jet production and EW effects compared to the theoretical cross section used in this analysis. The post-fit value of the nuisance parameter associated with the systematic uncertainty in the \( ttW \) background with \( N_{\text{jets}} = 7 \) is \( 0.18^{+0.73}_{-0.61} \). This corresponds to a 22% increase in the number of \( ttW \) events with seven jets. The post-fit value of the nuisance parameter for the systematic uncertainty of the \( ttW \)+jets background with \( N_{\text{jets}} \geq 8 \) is \( 0.25^{+0.56}_{-0.42} \), corresponding to a 65% increase in the number of \( ttW \) events with eight or more jets. As a result of these increases and of the change in the \( ttW \) background normalisation factor \( NF_{ttW} \), the overall \( ttW \) background yield in the signal-enriched region with a BDT score above zero increased from the 12.4 \( \pm \) 8.8 events predicted to 23.2 \( \pm \) 10.1 events after the fit to data. Apart from the uncertainties discussed above, no other nuisance parameters are found to be significantly adjusted or constrained by the fit.

Figure 4 shows the yields or the discriminating variable distributions used in the fit in each CR. Good agreement is observed between data and the post-fit computation.

In order to check the \( ttW \)+jets background normalisation and modelling, a validation region is defined, based on the fact that the \( ttW \)+jets process is charge asymmetric. The difference between the number of events with a positive sum and the number of events with a negative sum of the charges of the selected leptons. The band includes the total uncertainty of the post-fit computation. The ratio of the data to the total post-fit computation is shown in the lower panel. The first and last bins contain underflow and overflow events, respectively.

The distributions for some of the key analysis variables are shown in Fig. 6 for the events in the signal region and in Fig. 7 for events in a signal-enriched region with a BDT score above zero.

The uncertainties impacting \( \mu \) are summarised in Table 4. Apart from the theoretical uncertainty of the signal cross section, the largest systematic uncertainty comes from the modelling of the \( ttW \)+jets process. Within the uncertainties of the background modelling, the impact of the uncertainty in \( tt\bar{t} \) production is also significant. The expected cross section of this process is only of the order of 10% of \( \sigma_{ttH} \). However, the shape of the BDT score distribution for \( tt\bar{t} \) production is
Fig. 6 Post-fit comparison between data and prediction for signal region events for the distributions of: the sum of $b$-tagging pseudo-continuous scores of the jets in the event (top left), the minimum distance between two leptons among all possible pairs (top right), the multiplicity of jets (bottom left) and the multiplicity of $b$-tag jets (bottom right). The band includes the total uncertainty of the post-fit computation. The ratio of the data to the total post-fit computation is shown in the lower panel. The first and last bins contain underflow and overflow events, respectively. See Sect. 5 for the definitions of the different background categories.

The stability of the result has been checked. The fit was repeated with the data split according to year or by splitting the signal region into two regions with either same-sign dilepton events or events with at least three leptons. Differ-

similar to the one for signal, which leads to this uncertainty having a sizeable impact on $\mu$. In order to test the sensitivity of the $t\bar{t}t\bar{t}$ measurement to the value of the $t\bar{t}$ cross section, the fit was also performed assuming a $t\bar{t}$ cross section five times larger than the SM cross section. In that case the $t\bar{t}t\bar{t}$ signal strength $\mu$ decreases by 10% while the fitted normalisation factors are mostly unaffected.

The stability of the result has been checked. The fit was repeated with the data split according to year or by splitting the signal region into two regions with either same-sign dilepton events or events with at least three leptons. Differ-


Fig. 7 Post-fit comparison between data and prediction for signal region events with a BDT score greater than zero for the distributions of: the sum of $b$-tagging pseudo-continuous scores of the jets in the event (top left), the minimum distance between two leptons among all possible pairs (top right), the multiplicity of jets (bottom left) and the multiplicity of $b$-tag jets (bottom right). The band includes the total uncertainty of the post-fit computation. The ratio of the data to the total post-fit computation is shown in the lower panel. The first and last bins contain underflow and overflow events, respectively. See Sect. 5 for the definitions of the different background categories.

Different fits were also performed by using only positively charged same-sign lepton pairs or only negatively charged same-sign lepton pairs. All these tests showed compatible $\mu$ values.

An additional test was performed by splitting the SR into five regions according to the number of leptons and $b$-tagged jets and by fitting the $H_T$ distribution in each region. The BDT score is therefore not used in this test. The observed (expected) significance is found to be 4.3 (2.1) and the fitted signal strength is $2.2^{+0.9}_{-0.6}$. This result is consistent with the result from the default fit.

9 Conclusion

A search is presented for four-top-quark production using an integrated luminosity of 139 fb$^{-1}$ of proton–proton collision data at $\sqrt{s} = 13$ TeV collected by the ATLAS detector at the
LHC. The observed (expected) significance with respect to the background-only hypothesis is 4.3 (2.4) standard deviations, providing evidence for this process assuming the Standard Model four-top-quark production properties. The four-top-quark production cross section is measured to be $24^{+7}_{-6}$ fb which is consistent within 1.7 standard deviations with the Standard Model expectation of $\sigma_{\bar{t}t\bar{t}t} = 12.0 \pm 2.4$ fb.

Acknowledgements  We thank CERN for the very successful operation of the LHC, as well as the support staff from our institutions without whom ATLAS could not have been operated efficiently. We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BMWFW and FWF, Austria; ANAS, Azerbaijan; SSTC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN; ANID, Chile; CAS, MOST and NSFC, China; COLCIENCIAS, Colombia; MSMT CR, MPO CR and VSC CR, Czech Republic; DNRF and DNSRC, Denmark; IN2P3-CNRS and CEA-DRF/IRFU, France; SRNSFG, Georgia; BMBF, HGF and MPG, Germany; GSRT, Greece; RGC and Hong Kong SAR, China; ISF and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRSST, Morocco; NWO, Netherlands; RCN, Norway; MINSW and NCN, Poland; FCT, Portugal; MNE/IFA, Romania; MES of Russia and NRC KI, Russia Federation; JINR; MESTD, Serbia; MSSSR, Slovakia; ARRS and MIZŠ, Slovenia; DST/NRF, South Africa; MICINN, Spain; SRC and Wallenberg Foundation, Sweden; SERI, SNSF and Cantons of Bern and Geneva, Switzerland; MOST, Taiwan; TAEK, Turkey; STFC, UK; DOE and NSF, USA. In addition, individual groups and members have received support from BCKDF, CANARIE, Compute Canada and CRC, Canada; ERC, ERDF, Horizon 2020, Marie Skłodowska-Curie Actions and COST, European Union; Investissements d’Avenir Labex, Investissements d’Avenir Idex and ANR, France; DFG and AvH Foundation, Germany; Herakleitos, Thales and Aristeia programmes co-financed by EU-ESF and the Greek NSRF; Greece; BSF-NSF and GIF, Israel; La Caixa Banking Foundation, CERCA Programme Generalitat de Catalunya and PROMETEO and GenT Programmes Generalitat Valenciana, Spain; Góran Gustafssons Stiftelse, Sweden; The Royal Society and Leverhulme Trust, UK. The crucial computing support from all WLCG partners is acknowledged gratefully, in particular from CERN, the ATLAS Tier-1 facilities at TRIUMF (Canada), NDGF (Denmark, Norway, Sweden), CC-IN2P3 (France), KIT/GridKA (Germany), INFN-CNAF (Italy), NL-T1 (Netherlands), PIC (Spain), ASGC (Taiwan), RAL (UK) and BNL (USA), the Tier-2 facilities worldwide and large non-WLCG resource providers. Major contributors of computing resources are listed in Ref. [82].

Data Availability Statement This manuscript has no associated data or the data will not be deposited. [Authors’ comment: All ATLAS scientific output is published in journals, and preliminary results are made available in Conference Notes. All are openly available, without restriction on use by external parties beyond copyright law and the standard conditions agreed by CERN. Data associated with journal publications are also made available: tables and data from plots (e.g. cross section values, likelihood profiles, selection efficiencies, cross section limits, ...) are stored in appropriate repositories such as HEPDATA (http://hepdata.cedar.ac.uk/). ATLAS also strives to make additional material related to the paper available that allows a reinterpretation of the data in the context of new theoretical models. For example, an extended encapsulation of the analysis is often provided for measurements in the framework of RIVET (http://rivet.hepforge.org/). This information is taken from the ATLAS Data Access Policy, which is a public document that can be downloaded from http://opendata.cern.ch/record/413 [opendata.cern.ch].]

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Funded by SCOAP3.

References

1. Q.-H. Cao, S.-L. Chen, Y. Liu, Probing Higgs width and top quark Yukawa coupling from $t\bar{t}H$ and $t\bar{t}t\bar{t}$ productions. Phys. Rev. D.


ATLAS Collaboration

G. Aad\textsuperscript{102}, B. Abbott\textsuperscript{128}, D. C. Abbott\textsuperscript{103}, A. Abed Abud\textsuperscript{36}, K. Abeling\textsuperscript{53}, D. K. Abhayasinghe\textsuperscript{94}, S. H. Abidjani\textsuperscript{161}, O. S. AbouZeid\textsuperscript{40}, N. L. Abraham\textsuperscript{155}, H. Abramowicz\textsuperscript{160}, H. Abreu\textsuperscript{159}, Y. Abubakr\textsuperscript{1}, B. S. Acharya\textsuperscript{67a,67b,n}, B. Achkar\textsuperscript{51}, L. Adam\textsuperscript{100}, C. Adam Boundaries\textsuperscript{8}, L. Adamczyk\textsuperscript{84a}, L. Adamek\textsuperscript{166}, J. Adelman\textsuperscript{121}, M. Adersberger\textsuperscript{114}, A. Adiguzel\textsuperscript{12c}, N. L. Abraham\textsuperscript{155}, H. Abramowicz\textsuperscript{160}, D. Barberis\textsuperscript{55a,55b}, H. Bahrasemani\textsuperscript{151}, A. J. Bailey\textsuperscript{173}, V. Andrei\textsuperscript{61a}, C. R. Anelli\textsuperscript{175}


1 Department of Physics, University of Adelaide, Adelaide, Australia
2 Physics Department, SUNY Albany, Albany, NY, USA
3 Department of Physics, University of Alberta, Edmonton, AB, Canada
4 (a)Department of Physics, Ankara University, Ankara, Turkey; (b)Application and Research Center for Advanced Studies, Istanbul Aydin University, Istanbul, Turkey; (c)Division of Physics, TOBB University of Economics and Technology, Ankara, Turkey
5 LAPP, Université Grenoble Alpes, Université Savoie Mont Blanc, CNRS/IN2P3, Annecy, France
6 High Energy Physics Division, Argonne National Laboratory, Argonne, IL, USA
7 Department of Physics, University of Arizona, Tucson, AZ, USA
8 Department of Physics, University of Texas at Arlington, Arlington, TX, USA
9 Physics Department, National and Kapodistrian University of Athens, Athens, Greece
10 Physics Department, National Technical University of Athens, Zografou, Greece
11 Department of Physics, University of Texas at Austin, Austin, TX, USA
12 (a)Faculty of Engineering and Natural Sciences, Bahcesehir University, Istanbul, Turkey; (b)Faculty of Engineering and Natural Sciences, Istanbul Bilgi University, Istanbul, Turkey; (c)Department of Physics, Bogazici University, Istanbul, Turkey; (d)Department of Physics Engineering, Gaziantep University, Gaziantep, Turkey
13 Institute of Physics, Azerbaijan Academy of Sciences, Baku, Azerbaijan
14 Institut de Física d’Altes Energies (IFAE), Barcelona Institute of Science and Technology, Barcelona, Spain
15 (a)Institute of High Energy Physics, Chinese Academy of Sciences, Beijing, China; (b)Physics Department, Tsinghua University, Beijing, China; (c)Department of Physics, Nanjing University, Nanjing, China; (d)University of Chinese Academy of Science (UCAS), Beijing, China
16 Institute of Physics, University of Belgrade, Belgrade, Serbia
17 Department for Physics and Technology, University of Bergen, Bergen, Norway
18 Physics Division, Lawrence Berkeley National Laboratory and University of California, Berkeley, CA, USA
19 Institut für Physik, Humboldt Universität zu Berlin, Berlin, Germany
20 Albert Einstein Center for Fundamental Physics and Laboratory for High Energy Physics, University of Bern, Bern, Switzerland
21 School of Physics and Astronomy, University of Birmingham, Birmingham, UK
22 (a)Facultad de Ciencias y Centro de Investigaciones, Universidad Antonio Nariño, Bogotá, Colombia; (b)Departamento de Física, Universidad Nacional de Colombia, Bogotá, Colombia
23 (a)Dipartimento di Fisica, INFN Bologna and Universita’ di Bologna, Bologna, Italy; (b)INFN Sezione di Bologna, Bologna, Italy
24 Physikalisches Institut, Universität Bonn, Bonn, Germany
25 Department of Physics, Boston University, Boston, MA, USA
26 Department of Physics, Brandeis University, Waltham, MA, USA
27 (a)Transilvania University of Brasov, Brasov, Romania; (b)Horia Hulubei National Institute of Physics and Nuclear Engineering, Bucharest, Romania; (c)Department of Physics, Alexandru Ioan Cuza University of Iasi, Iasi, Romania; (d)Physics Department, National Institute for Research and Development of Isotopic and Molecular Technologies, Cluj-Napoca, Romania; (e)University Politehnica Bucharest, Bucharest, Romania; (f)West University in Timisoara, Timisoara, Romania
28 (a)Faculty of Mathematics, Physics and Informatics, Comenius University, Bratislava, Slovak Republic; (b)Department of Subnuclear Physics, Institute of Experimental Physics of the Slovak Academy of Sciences, Kosice, Slovak Republic
29 Physics Department, Brookhaven National Laboratory, Upton, NY, USA
30 Departamento de Física, Universidad de Buenos Aires, Buenos Aires, Argentina
31 California State University, Long Beach, CA, USA
32 Cavendish Laboratory, University of Cambridge, Cambridge, UK
33 (a)Department of Physics, University of Cape Town, Cape Town, South Africa; (b)iThemba Labs, Cape Town, Western Cape, South Africa; (c)Department of Mechanical Engineering Science, University of Johannesburg, Johannesburg,
160 Raymond and Beverly Sackler School of Physics and Astronomy, Tel Aviv University, Tel Aviv, Israel
161 Department of Physics, Aristotle University of Thessaloniki, Thessaloniki, Greece
162 International Center for Elementary Particle Physics and Department of Physics, University of Tokyo, Tokyo, Japan
163 Graduate School of Science and Technology, Tokyo Metropolitan University, Tokyo, Japan
164 Department of Physics, Tokyo Institute of Technology, Tokyo, Japan
165 Tomsk State University, Tomsk, Russia
166 Department of Physics, University of Toronto, Toronto, ON, Canada
167 Division of Physics and Tomonaga Center for the History of the Universe, Faculty of Pure and Applied Sciences, University of Tsukuba, Tsukuba, Japan
168 Department of Physics and Astronomy, Tufts University, Medford, MA, USA
169 Department of Physics and Astronomy, University of California Irvine, Irvine, CA, USA
170 Department of Physics and Astronomy, University of Uppsala, Uppsala, Sweden
171 Department of Physics, University of Illinois, Urbana, IL, USA
172 Instituto de Física Corpuscular (IFIC), Centro Mixto Universidad de Valencia-CSIC, Valencia, Spain
173 Department of Physics, University of British Columbia, Vancouver, BC, Canada
174 Department of Physics and Astronomy, University of Victoria, Victoria, BC, Canada
175 Fakultät für Physik und Astronomie, Julius-Maximilians-Universität Würzburg, Würzburg, Germany
176 Department of Physics, University of Warwick, Coventry, UK
177 Waseda University, Tokyo, Japan
178 Department of Particle Physics and Astrophysics, Weizmann Institute of Science, Rehovot, Israel
179 Department of Physics, University of Wisconsin, Madison, WI, USA
180 Fakultät für Mathematik und Naturwissenschaften, Fachgruppe Physik, Bergische Universität Wuppertal, Wuppertal, Germany
181 Department of Physics, Yale University, New Haven, CT, USA

a Also at Borough of Manhattan Community College, City University of New York, New York, NY, USA
b Also at Centro Studi e Ricerche Enrico Fermi, Italy
c Also at CERN, Geneva, Switzerland
d Also at CPPM, Aix-Marseille Université, CNRS/IN2P3, Marseille, France
e Also at Département de Physique Nucléaire et Corpusculaire, Université de Genève, Geneva, Switzerland
f Also at Departamento de Física de la Universitat Autonoma de Barcelona, Barcelona, Spain
g Also at Department of Financial and Management Engineering, University of the Aegean, Chios, Greece
h Also at Department of Physics and Astronomy, Michigan State University, East Lansing, MI, USA
i Also at Department of Physics and Astronomy, University of Louisville, Louisville, KY, USA
j Also at Department of Physics, Ben Gurion University of the Negev, Beersheba, Israel
k Also at Department of Physics, California State University, East Bay, USA
l Also at Department of Physics, California State University, Sacramento, USA
m Also at Department of Physics, King’s College London, London, UK
n Also at Department of Physics, St. Petersburg State Polytechnical University, Saint Petersburg, Russia
o Also at Department of Physics, University of Fribourg, Fribourg, Switzerland
p Also at Dipartimento di Matematica, Informatica e Fisica, Università di Udine, Udine, Italy
q Also at Dipartimento di Matematica, Informatica e Fisica, Università di Udine, Udine, Italy
r Also at Faculty of Physics, M.V. Lomonosov Moscow State University, Moscow, Russia
s Also at Faculty of Engineering, Giresun University, Giresun, Turkey
t Also at Graduate School of Science, Osaka University, Osaka, Japan
u Also at Hellenic Open University, Patras, Greece
v Also at IJCLab, Université Paris-Saclay, CNRS/IN2P3, 91405, Orsay, France
w Also at Institut Catalana de Recerca i Estudis Avancats, ICREA, Barcelona, Spain
x Also at Institut für Experimentalphysik, Universität Hamburg, Hamburg, Germany
y Also at Institute for Mathematics, Astrophysics and Particle Physics, Radboud University Nijmegen/Nikhef, Nijmegen, The Netherlands
Also at Institute for Nuclear Research and Nuclear Energy (INRNE) of the Bulgarian Academy of Sciences, Sofia, Bulgaria
Also at Institute for Particle and Nuclear Physics, Wigner Research Centre for Physics, Budapest, Hungary
Also at Institute of Particle Physics (IPP), Canada
Also at Institute of Physics, Azerbaijan Academy of Sciences, Baku, Azerbaijan
Also at Instituto de Fisica Teorica, IFT-UAM/CSIC, Madrid, Spain
Also at Joint Institute for Nuclear Research, Dubna, Russia
Also at Louisiana Tech University, Ruston, LA, USA
Also at Moscow Institute of Physics and Technology State University, Dolgoprudny, Russia
Also at National Research Nuclear University MEPhI, Moscow, Russia
Also at Physics Department, An-Najah National University, Nablus, Palestine
Also at Physikalisches Institut, Albert-Ludwigs-Universität Freiburg, Freiburg, Germany
Also at The City College of New York, New York, NY, USA
Also at TRIUMF, Vancouver, BC, Canada
Also at Universita di Napoli Parthenope, Naples, Italy
Also at University of Chinese Academy of Sciences (UCAS), Beijing, China
*Deceased